



Australian Bureau of Statistics

1351.0.55.034 - Research Paper: Sampling-Based Clerical Review Methods in Probabilistic Linking, May 2011

Latest ISSUE Released at 11:30 AM (CANBERRA TIME) 25/05/2011 First Issue

Summary

About this Release

Probabilistic data linking aims to link records that are believed to belong to the same person from two different data sets. Candidate record pairs are given a weight based on the degree of agreement between fields on the two records. Record pairs with a weight above some upper cut-off are declared links while those with a weight below some lower cut-off are declared non-links. However, there are many record pairs that cannot be automatically assigned a status and are designated for clerical review. Clerical review is a time-consuming and resource-intensive stage of the data linking process requiring careful visual inspection and keyboard use. Acceptance sampling is proposed to dramatically reduce the amount of clerical inspection. Sampling is also proposed as a method to provide an accurate and reliable means of assessing and setting the most appropriate clerical review bounds.

© Commonwealth of Australia

All data and other material produced by the Australian Bureau of Statistics (ABS) constitutes Commonwealth copyright administered by the ABS. The ABS reserves the right to set out the terms and conditions for the use of such material. Unless otherwise noted, all material on this website – except the ABS logo, the Commonwealth Coat of Arms, and any material protected by a trade mark – is licensed under a Creative Commons Attribution 2.5 Australia licence